

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Tryggvason, Karl  
Kallunki, Pekka  
Pyke, Charles

(ii) TITLE OF INVENTION: Laminin Chains: Diagnostic Uses

(iii) NUMBER OF SEQUENCES: 20

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fay Sharpe Fagan Minnich & McKee  
(B) STREET: 1100 Superior Ave, Suite 700  
(C) CITY: Cleveland  
(D) STATE: Ohio  
(E) COUNTRY: USA  
(F) ZIP: 44114

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Unknown  
(B) FILING DATE: 08 January 2001  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/663,147  
(B) FILING DATE: 150-September 2000  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Minnich, Richard, J.  
(B) REGISTRATION NUMBER: 24,175  
(C) REFERENCE/DOCKET NUMBER: TRV 20014

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 216-861-5582  
(B) TELEFAX: 216-241-1666

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligomer primers"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCTCACCAA GACTTACACA

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAATCACTGA GCAGCTGAAC

20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAGTACCAGA ACCGAGTTTG

20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTGGTTACCA GGCTTGAGAG

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTACTGCGGA ATCTCACAGC

20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TACACTGTTC AACCCAGGGT

20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAACAAGCCC TCTCACTGGT

20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCGGAGACTG TGCTGATAAG

20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CATACCTCTC TACATGGCAT

20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGTCTCGCTG AATCTCTCTT

20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTACAACTAG CATGGTGCCC

20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 118..183

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 118..3699

(ix) FEATURE:

- (A) NAME/KEY: polyA\_site
- (B) LOCATION: 4433

(ix) FEATURE:

- (A) NAME/KEY: polyA\_site
- (B) LOCATION: 5195

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GenBank accession number: U00166.1 (H. pylori strain NCTC 11637 genome) (NCBI)

GACCACCTGA TCGAAGGAAA AGGAAGGCAC AGCGGAGCGC AGAGTGAGAA CCACCAACCG	60
AGGCGCCGGG CAGCGACCCC TGCAGCGGAG ACAGAGACTG AGCGGCCCCG CACCGCC	117
ATG CCT GCG CTC TGG CTG GGC TGC TGC CTC TGC TTC TCG CTC CTC CTG Met Pro Ala Leu Trp Leu Gly Cys Cys Leu Cys Phe Ser Leu Leu Leu 1 5 10 15	165
CCC GCA GCC CGG GCC ACC TCC AGG AGG GAA GTC TGT GAT TGC AAT GGG Pro Ala Ala Arg Ala Thr Ser Arg Arg Glu Val Cys Asp Cys Asn Gly 20 25 30	213
AAG TCC AGG CAG TGT ATC TTT GAT CGG GAA CTT CAC AGA CAA ACT GGT Lys Ser Arg Gln Cys Ile Phe Asp Arg Glu Leu His Arg Gln Thr Gly 35 40 45	261
AAT GGA TTC CGC TGC CTC AAC TGC AAT GAC AAC ACT GAT GGC ATT CAC Asn Gly Phe Arg Cys Leu Asn Cys Asn Asp Asn Thr Asp Gly Ile His 50 55 60	309
TGC GAG AAG TGC AAG AAT GGC TTT TAC CGG CAC AGA GAA AGG GAC CGC Cys Glu Lys Cys Lys Asn Gly Phe Tyr Arg His Arg Glu Arg Asp Arg 65 70 75 80	357
TGT TTG CCC TGC AAT TGT AAC TCC AAA GGT TCT CTT AGT GCT CGA TGT Cys Leu Pro Cys Asn Cys Asn Ser Lys Gly Ser Leu Ser Ala Arg Cys 85 90 95	405
GAC AAC TCT GGA CGG TGC AGC TGT AAA CCA GGT GTG ACA GGA GCC AGA Asp Asn Ser Gly Arg Cys Ser Cys Lys Pro Gly Val Thr Gly Ala Arg 100 105 110	453
TGC GAC CGA TGT CTG CCA GGC TTC CAC ATG CTC ACG GAT GCG GGG TGC Cys Asp Arg Cys Leu Pro Gly Phe His Met Leu Thr Asp Ala Gly Cys 115 120 125	501
ACC CAA GAC CAG AGA CTG CTA GAC TCC AAG TGT GAC TGT GAC CCA GCT Thr Gln Asp Gln Arg Leu Leu Asp Ser Lys Cys Asp Cys Asp Pro Ala 130 135 140	549
GGC ATC GCA GGG CCC TGT GAC GCG GGC CGC TGT GTC TGC AAG CCA GCT Gly Ile Ala Gly Pro Cys Asp Ala Gly Arg Cys Val Cys Lys Pro Ala 145 150 155 160	597
GTT ACT GGA GAA CGC TGT GAT AGG TGT CGA TCA GGT TAC TAT AAT CTG Val Thr Gly Glu Arg Cys Asp Arg Cys Arg Ser Gly Tyr Tyr Asn Leu 165 170 175	645
GAT GGG GGG AAC CCT GAG GGC TGT ACC CAG TGT TTC TGC TAT GGG CAT Asp Gly Gly Asn Pro Glu Gly Cys Thr Gln Cys Phe Cys Tyr Gly His 180 185 190	693
TCA GCC AGC TGC CGC AGC TCT GCA GAA TAC AGT GTC CAT AAG ATC ACC Ser Ala Ser Cys Arg Ser Ser Ala Glu Tyr Ser Val His Lys Ile Thr 195 200 205	741
TCT ACC TTT CAT CAA GAT GTT GAT GGC TGG AAG GCT GTC CAA CGA AAT Ser Thr Phe His Gln Asp Val Asp Gly Trp Lys Ala Val Gln Arg Asn	789

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

210	215	220	
GGG TCT CCT GCA AAG CTC CAA TGG TCA CAG CGC CAT CAA GAT GTG TTT Gly Ser Pro Ala Lys Leu Gln Trp Ser Gln Arg His Gln Asp Val Phe 225 230 235 240	837		
AGC TCA GCC CAA CGA CTA GAT CCT GTC TAT TTT GTG GCT CCT GCC AAA Ser Ser Ala Gln Arg Leu Asp Pro Val Tyr Phe Val Ala Pro Ala Lys 245 250 255	885		
TTT CTT GGG AAT CAA CAG GTG AGC TAT GGG CAA AGC CTG TCC TTT GAC Phe Leu Gly Asn Gln Gln Val Ser Tyr Gly Gln Ser Leu Ser Phe Asp 260 265 270	933		
TAC CGT GTG GAC AGA GGA GGC AGA CAC CCA TCT GCC CAT GAT GTG ATC Tyr Arg Val Asp Arg Gly Gly Arg His Pro Ser Ala His Asp Val Ile 275 280 285	981		
CTG GAA GGT GCT GGT CTA CGG ATC ACA GCT CCC TTG ATG CCA CTT GGC Leu Glu Gly Ala Gly Leu Arg Ile Thr Ala Pro Leu Met Pro Leu Gly 290 295 300	1029		
AAG ACA CTG CCT TGT GGG CTC ACC AAG ACT TAC ACA TTC AGG TTA AAT Lys Thr Leu Pro Cys Gly Leu Thr Lys Thr Tyr Thr Phe Arg Leu Asn 305 310 315 320	1077		
GAG CAT CCA AGC AAT AAT TGG AGC CCC CAG CTG AGT TAC TTT GAG TAT Glu His Pro Ser Asn Asn Trp Ser Pro Gln Leu Ser Tyr Phe Glu Tyr 325 330 335	1125		
CGA AGG TTA CTG CGG AAT CTC ACA GCC CTC CGC ATC CGA GCT ACA TAT Arg Arg Leu Leu Arg Asn Leu Thr Ala Leu Arg Ile Arg Ala Thr Tyr 340 345 350	1173		
GGA GAA TAC AGT ACT GGG TAC ATT GAC AAT GTG ACC CTG ATT TCA GCC Gly Glu Tyr Ser Thr Gly Tyr Ile Asp Asn Val Thr Leu Ile Ser Ala 355 360 365	1221		
CGC CCT GTC TCT GGA GCC CCA GCA CCC TGG GTT GAA CAG TGT ATA TGT Arg Pro Val Ser Gly Ala Pro Ala Pro Trp Val Glu Gln Cys Ile Cys 370 375 380	1269		
CCT GTT GGG TAC AAG GGG CAA TTC TGC CAG GAT TGT GCT TCT GGC TAC Pro Val Gly Tyr Lys Gly Gln Phe Cys Gln Asp Cys Ala Ser Gly Tyr 385 390 395 400	1317		
AAG AGA GAT TCA GCG AGA CTG GGG CCT TTT GGC ACC TGT ATT CCT TGT Lys Arg Asp Ser Ala Arg Leu Gly Pro Phe Gly Thr Cys Ile Pro Cys 405 410 415	1365		
AAC TGT CAA GGG GGA GGG GCC TGT GAT CCA GAC ACA GGA GAT TGT TAT Asn Cys Gln Gly Gly Gly Ala Cys Asp Pro Asp Thr Gly Asp Cys Tyr 420 425 430	1413		
TCA GGG GAT GAG AAT CCT GAC ATT GAG TGT GCT GAC TGC CCA ATT GGT Ser Gly Asp Glu Asn Pro Asp Ile Glu Cys Ala Asp Cys Pro Ile Gly 435 440 445	1461		
TTC TAC AAC GAT CCG CAC GAC CCC CGC AGC TGC AAG CCA TGT CCC TGT	1509		



AAC Asn	AGC Ser 690	TAC Tyr	CAG Gln	AGC Ser	CGC Arg	CTG Leu 695	GAT Asp	GAC Asp	CTC Leu	AAG Lys	ATG Met 700	ACT Thr	GTG Val	GAA Glu	AGA Arg	2229
GTT Val 705	CGG Arg	GCT Ala	CTG Leu	GGA Gly	AGT Ser 710	CAG Gln	TAC Tyr	CAG Gln	AAC Asn	CGA Arg 715	GTT Val	CGG Arg	GAT Asp	ACT Thr	CAC His 720	2277
AGG Arg	CTC Leu	ATC Ile	ACT Thr	CAG Gln 725	ATG Met	CAG Gln	CTG Leu	AGC Ser	CTG Leu 730	GCA Ala	GAA Glu	AGT Ser	GAA Glu	GCT Ala 735	TCC Ser	2325
TTG Leu	GGA Gly	AAC Asn	ACT Thr 740	AAC Asn	ATT Ile	CCT Pro	GCC Ala	TCA Ser 745	GAC Asp	CAC His	TAC Tyr	GTG Val 750	GGG Gly	CCA Pro	AAT Asn	2373
GGC Gly	TTT Phe 755	AAA Lys	AGT Ser	CTG Leu	GCT Ala	CAG Gln	GAG Glu 760	GCC Ala	ACA Thr	AGA Arg	TTA Leu 765	GCA Ala	GAA Glu	AGC Ser	CAC His	2421
GTT Val 770	GAG Glu	TCA Ser	GCC Ala	AGT Ser	AAC Asn	ATG Met 775	GAG Glu	CAA Gln	CTG Leu	ACA Thr 780	AGG Arg	GAA Glu	ACT Thr	GAG Glu	GAC Asp	2469
TAT Tyr 785	TCC Ser	AAA Lys	CAA Gln	GCC Ala	CTC Leu 790	TCA Ser	CTG Leu	GTG Val	CGC Arg 795	AAG Lys	GCC Ala	CTG Leu	CAT His	GAA Glu	GGA Gly 800	2517
GTC Val	GGA Gly	AGC Ser	GGA Gly	AGC Ser 805	GGT Gly	AGC Ser	CCG Pro	GAC Asp	GGT Gly 810	GCT Ala	GTG Val	GTG Val	CAA Gln	GGG Gly 815	CTT Leu	2565
GTG Val	GAA Glu	AAA Lys	TTG Leu 820	GAG Glu	AAA Lys	ACC Thr	AAG Lys	TCC Ser 825	CTG Leu	GCC Ala	CAG Gln	CAG Gln	TTG Leu 830	ACA Thr	AGG Arg	2613
GAG Glu	GCC Ala 835	ACT Thr	CAA Gln	GCG Ala	GAA Glu	ATT Ile	GAA Glu 840	GCA Ala	GAT Asp	AGG Arg	TCT Ser 845	TAT Tyr	CAG Gln	CAC His	AGT Ser	2661
CTC Leu	CGC Arg 850	CTC Leu	CTG Leu	GAT Asp	TCA Ser	GTG Val 855	TCT Ser	CCG Pro	CTT Leu	CAG Gln 860	GGA Gly	GTC Val	AGT Ser	GAT Asp	CAG Gln	2709
TCC Ser 865	TTT Phe	CAG Gln	GTG Val	GAA Glu 870	GAA Glu	GCA Ala	AAG Lys	AGG Arg	ATC Ile	AAA Lys 875	CAA Gln	AAA Lys	GCG Ala	GAT Asp	TCA Ser 880	2757
CTC Leu	TCA Ser	AGC Ser	CTG Leu	GTA Val 885	ACC Thr	AGG Arg	CAT His	ATG Met	GAT Asp 890	GAG Glu	TTC Phe	AAG Lys	CGT Arg	ACA Thr 895	CAA Gln	2805
AAG Lys	AAT Asn	CTG Leu 900	GGA Gly	AAC Asn	TGG Trp	AAA Lys	GAA Glu 905	GAA Glu	GCA Ala	CAG Gln	CAG Gln	CTC Leu 910	TTA Leu 910	CAG Gln	AAT Asn	2853
GGA Gly	AAA Lys	AGT Ser	GGG Gly	AGA Arg	GAG Glu	AAA Lys	TCA Ser	GAT Asp	CAG Gln	CTG Leu	CTT Leu	TCC Ser	CGT Arg	GCC Ala	AAT Asn	2901



1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150

	915		920		925	
CTT GCT AAA AGC AGA GCA CAA GAA GCA CTG AGT ATG GGC AAT GCC ACT						2949
Leu Ala Lys Ser Arg Ala Gln Glu Ala Leu Ser Met Gly Asn Ala Thr						
930			935		940	
TTT TAT GAA GTT GAG AGC ATC CTT AAA AAC CTC AGA GAG TTT GAC CTG						2997
Phe Tyr Glu Val Glu Ser Ile Leu Lys Asn Leu Arg Glu Phe Asp Leu						
945			950		955	960
CAG GTG GAC AAC AGA AAA GCA GAA GCT GAA GAA GCC ATG AAG AGA CTC						3045
Gln Val Asp Asn Arg Lys Ala Glu Ala Glu Glu Ala Met Lys Arg Leu						
965			970		975	
TCC TAC ATC AGC CAG AAG GTT TCA GAT GCC AGT GAC AAG ACC CAG CAA						3093
Ser Tyr Ile Ser Gln Lys Val Ser Asp Ala Ser Asp Lys Thr Gln Gln						
980			985		990	
GCA GAA AGA GCC CTG GGG AGC GCT GCT GCT GAT GCA CAG AGG GCA AAG						3141
Ala Glu Arg Ala Leu Gly Ser Ala Ala Ala Asp Ala Gln Arg Ala Lys						
995			1000		1005	
AAT GGG GCC GGG GAG GCC CTG GAA ATC TCC AGT GAG ATT GAA CAG GAG						3189
Asn Gly Ala Gly Glu Ala Leu Glu Ile Ser Ser Glu Ile Glu Gln Glu						
1010			1015		1020	
ATT GGG AGT CTG AAC TTG GAA GCC AAT GTG ACA GCA GAT GGA GCC TTG						3237
Ile Gly Ser Leu Asn Leu Glu Ala Asn Val Thr Ala Asp Gly Ala Leu						
1025			1030		1035	1040
GCC ATG GAA AAG GGA CTG GCC TCT CTG AAG AGT GAG ATG AGG GAA GTG						3285
Ala Met Glu Lys Gly Leu Ala Ser Leu Lys Ser Glu Met Arg Glu Val						
1045			1050		1055	
GAA GGA GAG CTG GAA AGG AAG GAG CTG GAG TTT GAC ACG AAT ATG GAT						3333
Glu Gly Glu Leu Glu Arg Lys Glu Leu Glu Phe Asp Thr Asn Met Asp						
1060			1065		1070	
GCA GTA CAG ATG GTG ATT ACA GAA GCC CAG AAG GTT GAT ACC AGA GCC						3381
Ala Val Gln Met Val Ile Thr Glu Ala Gln Lys Val Asp Thr Arg Ala						
1075			1080		1085	
AAG AAC GCT GGG GTT ACA ATC CAA GAC ACA CTC AAC ACA TTA GAC GGC						3429
Lys Asn Ala Gly Val Thr Ile Gln Asp Thr Leu Asn Thr Leu Asp Gly						
1090			1095		1100	
CTC CTG CAT CTG ATG GAC CAG CCT CTC AGT GTA GAT GAA GAG GGG CTG						3477
Leu Leu His Leu Met Asp Gln Pro Leu Ser Val Asp Glu Glu Gly Leu						
1105			1110		1115	1120
GTC TTA CTG GAG CAG AAG CTT TCC CGA GCC AAG ACC CAG ATC AAC AGC						3525
Val Leu Leu Glu Gln Lys Leu Ser Arg Ala Lys Thr Gln Ile Asn Ser						
1125			1130		1135	
CAA CTG CGG CCC ATG ATG TCA GAG CTG GAA GAG AGG GCA CGT CAG CAG						3573
Gln Leu Arg Pro Met Met Ser Glu Leu Glu Glu Arg Ala Arg Gln Gln						
1140			1145		1150	



## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1194 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

Met  Pro  Ala  Leu  Trp  Leu  Gly  Cys  Cys  Leu  Cys  Phe  Ser  Leu  Leu  Leu
  1      5      10      15

Pro  Ala  Ala  Arg  Ala  Thr  Ser  Arg  Arg  Glu  Val  Cys  Asp  Cys  Asn  Gly
      20      25      30

Lys  Ser  Arg  Gln  Cys  Ile  Phe  Asp  Arg  Glu  Leu  His  Arg  Gln  Thr  Gly
      35      40      45

Asn  Gly  Phe  Arg  Cys  Leu  Asn  Cys  Asn  Asp  Asn  Thr  Asp  Gly  Ile  His
      50      55      60

Cys  Glu  Lys  Cys  Lys  Asn  Gly  Phe  Tyr  Arg  His  Arg  Glu  Arg  Asp  Arg
      65      70      75      80

Cys  Leu  Pro  Cys  Asn  Cys  Asn  Ser  Lys  Gly  Ser  Leu  Ser  Ala  Arg  Cys
      85      90      95

Asp  Asn  Ser  Gly  Arg  Cys  Ser  Cys  Lys  Pro  Gly  Val  Thr  Gly  Ala  Arg
      100     105     110

Cys  Asp  Arg  Cys  Leu  Pro  Gly  Phe  His  Met  Leu  Thr  Asp  Ala  Gly  Cys
      115     120     125

Thr  Gln  Asp  Gln  Arg  Leu  Leu  Asp  Ser  Lys  Cys  Asp  Cys  Asp  Pro  Ala
      130     135     140

Gly  Ile  Ala  Gly  Pro  Cys  Asp  Ala  Gly  Arg  Cys  Val  Cys  Lys  Pro  Ala
      145     150     155     160

Val  Thr  Gly  Glu  Arg  Cys  Asp  Arg  Cys  Arg  Ser  Gly  Tyr  Tyr  Asn  Leu
      165     170     175

Asp  Gly  Gly  Asn  Pro  Glu  Gly  Cys  Thr  Gln  Cys  Phe  Cys  Tyr  Gly  His
      180     185     190

Ser  Ala  Ser  Cys  Arg  Ser  Ser  Ala  Glu  Tyr  Ser  Val  His  Lys  Ile  Thr
      195     200     205

Ser  Thr  Phe  His  Gln  Asp  Val  Asp  Gly  Trp  Lys  Ala  Val  Gln  Arg  Asn
      210     215     220

Gly  Ser  Pro  Ala  Lys  Leu  Gln  Trp  Ser  Gln  Arg  His  Gln  Asp  Val  Phe
      225     230     235     240

Ser  Ser  Ala  Gln  Arg  Leu  Asp  Pro  Val  Tyr  Phe  Val  Ala  Pro  Ala  Lys

```

245								250					255				
Phe	Leu	Gly	Asn	Gln	Gln	Val	Ser	Tyr	Gly	Gln	Ser	Leu	Ser	Phe	Asp		
			260				265						270				
Tyr	Arg	Val	Asp	Arg	Gly	Gly	Arg	His	Pro	Ser	Ala	His	Asp	Val	Ile		
			275				280						285				
Leu	Glu	Gly	Ala	Gly	Leu	Arg	Ile	Thr	Ala	Pro	Leu	Met	Pro	Leu	Gly		
			290				295						300				
Lys	Thr	Leu	Pro	Cys	Gly	Leu	Thr	Lys	Thr	Tyr	Thr	Phe	Arg	Leu	Asn		
305						310						320					
Glu	His	Pro	Ser	Asn	Asn	Trp	Ser	Pro	Gln	Leu	Ser	Tyr	Phe	Glu	Tyr		
			325						330			335					
Arg	Arg	Leu	Leu	Arg	Asn	Leu	Thr	Ala	Leu	Arg	Ile	Arg	Ala	Thr	Tyr		
			340						345			350					
Gly	Glu	Tyr	Ser	Thr	Gly	Tyr	Ile	Asp	Asn	Val	Thr	Leu	Ile	Ser	Ala		
			355						360			365					
Arg	Pro	Val	Ser	Gly	Ala	Pro	Ala	Pro	Trp	Val	Glu	Gln	Cys	Ile	Cys		
370						375						380					
Pro	Val	Gly	Tyr	Lys	Gly	Gln	Phe	Cys	Gln	Asp	Cys	Ala	Ser	Gly	Tyr		
385						390						400					
Lys	Arg	Asp	Ser	Ala	Arg	Leu	Gly	Pro	Phe	Gly	Thr	Cys	Ile	Pro	Cys		
			405						410			415					
Asn	Cys	Gln	Gly	Gly	Gly	Ala	Cys	Asp	Pro	Asp	Thr	Gly	Asp	Cys	Tyr		
			420						425			430					
Ser	Gly	Asp	Glu	Asn	Pro	Asp	Ile	Glu	Cys	Ala	Asp	Cys	Pro	Ile	Gly		
			435						440			445					
Phe	Tyr	Asn	Asp	Pro	His	Asp	Pro	Arg	Ser	Cys	Lys	Pro	Cys	Pro	Cys		
450						455						460					
His	Asn	Gly	Phe	Ser	Cys	Ser	Val	Ile	Pro	Glu	Thr	Glu	Glu	Val	Val		
465						470						475					
Cys	Asn	Asn	Cys	Pro	Pro	Gly	Val	Thr	Gly	Ala	Arg	Cys	Glu	Leu	Cys		
			485						490			495					
Ala	Asp	Gly	Tyr	Phe	Gly	Asp	Pro	Phe	Gly	Glu	His	Gly	Pro	Val	Arg		
			500						505			510					
Pro	Cys	Gln	Pro	Cys	Gln	Cys	Asn	Ser	Asn	Val	Asp	Pro	Ser	Ala	Ser		
			515						520			525					
Gly	Asn	Cys	Asp	Arg	Leu	Thr	Gly	Arg	Cys	Leu	Lys	Cys	Ile	His	Asn		
530						535						540					
Thr	Ala	Gly	Ile	Tyr	Cys	Asp	Gln	Cys	Lys	Ala	Gly	Tyr	Phe	Gly	Asp		
545						550						555					
Pro	Leu	Ala	Pro	Asn	Pro	Ala	Asp	Lys	Cys	Arg	Ala	Cys	Asn	Cys	Asn		

575

Pro	Met	Gly	Ser	Glu	Pro	Val	Gly	Cys	Arg	Ser	Asp	Gly	Thr	Cys	Val
			580				585						590		
Cys	Lys	Pro	Gly	Phe	Gly	Gly	Pro	Asn	Cys	Glu	His	Gly	Ala	Phe	Ser
			595				600						605		
Cys	Pro	Ala	Cys	Tyr	Asn	Gln	Val	Lys	Ile	Gln	Met	Asp	Gln	Phe	Met
			610				615						620		
Gln	Gln	Leu	Gln	Arg	Met	Glu	Ala	Leu	Ile	Ser	Lys	Ala	Gln	Gly	Gly
			625				630						635		
Asp	Gly	Val	Val	Pro	Asp	Thr	Glu	Leu	Glu	Gly	Arg	Met	Gln	Gln	Ala
			645				650						655		
Glu	Gln	Ala	Leu	Gln	Asp	Ile	Leu	Arg	Asp	Ala	Gln	Ile	Ser	Glu	Gly
			660				665						670		
Ala	Ser	Arg	Ser	Leu	Gly	Leu	Gln	Leu	Ala	Lys	Val	Arg	Ser	Gln	Glu
			675				680						685		
Asn	Ser	Tyr	Gln	Ser	Arg	Leu	Asp	Asp	Leu	Lys	Met	Thr	Val	Glu	Arg
			690				695						700		
Val	Arg	Ala	Leu	Gly	Ser	Gln	Tyr	Gln	Asn	Arg	Val	Arg	Asp	Thr	His
			705				710						715		
Arg	Leu	Ile	Thr	Gln	Met	Gln	Leu	Ser	Leu	Ala	Glu	Ser	Glu	Ala	Ser
			725				730						735		
Leu	Gly	Asn	Thr	Asn	Ile	Pro	Ala	Ser	Asp	His	Tyr	Val	Gly	Pro	Asn
			740				745						750		
Gly	Phe	Lys	Ser	Leu	Ala	Gln	Glu	Ala	Thr	Arg	Leu	Ala	Glu	Ser	His
			755				760						765		
Val	Glu	Ser	Ala	Ser	Asn	Met	Glu	Gln	Leu	Thr	Arg	Glu	Thr	Glu	Asp
			770				775						780		
Tyr	Ser	Lys	Gln	Ala	Leu	Ser	Leu	Val	Arg	Lys	Ala	Leu	His	Glu	Gly
			785				790						795		
Val	Gly	Ser	Gly	Ser	Gly	Ser	Pro	Asp	Gly	Ala	Val	Val	Gln	Gly	Leu
			805				810						815		
Val	Glu	Lys	Leu	Glu	Lys	Thr	Lys	Ser	Leu	Ala	Gln	Gln	Leu	Thr	Arg
			820				825						830		
Glu	Ala	Thr	Gln	Ala	Glu	Ile	Glu	Ala	Asp	Arg	Ser	Tyr	Gln	His	Ser
			835				840						845		
Leu	Arg	Leu	Leu	Asp	Ser	Val	Ser	Pro	Leu	Gln	Gly	Val	Ser	Asp	Gln
			850				855						860		
Ser	Phe	Gln	Val	Glu	Glu	Ala	Lys	Arg	Ile	Lys	Gln	Lys	Ala	Asp	Ser
			865				870						875		
Leu	Ser	Ser	Leu	Val	Thr	Arg	His	Met	Asp	Glu	Phe	Lys	Arg	Thr	Gln



- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4316 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 118..183

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 118..3453

(ix) FEATURE:

- (A) NAME/KEY: repeat\_unit
- (B) LOCATION: 4021..4316
- (D) OTHER INFORMATION: /rpt\_type= "other"

/rpt\_family= "HUMAN ALU"

(ix) FEATURE:

- (A) NAME/KEY: polyA\_site
- (B) LOCATION: 4296

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GACCACCTGA TCGAAGGAAA AGGAAGGCAC AGCGGAGCGC AGAGTGAGAA CCACCAACCG	60
AGGCGCCGGG CAGCGACCCC TGCAGCGGAG ACAGAGACTG AGCGGCCCGG CACCGCC	117
ATG CCT GCG CTC TGG CTG GGC TGC TGC CTC TGC TTC TCG CTC CTC CTG	165
Met Pro Ala Leu Trp Leu Gly Cys Cys Leu Cys Phe Ser Leu Leu Leu	
1195 1200 1205 1210	
CCC GCA GCC CGG GCC ACC TCC AGG AGG GAA GTC TGT GAT TGC AAT GGG	213
Pro Ala Ala Arg Ala Thr Ser Arg Arg Glu Val Cys Asp Cys Asn Gly	
1215 1220 1225	
AAG TCC AGG CAG TGT ATC TTT GAT CGG GAA CTT CAC AGA CAA ACT GGT	261
Lys Ser Arg Gln Cys Ile Phe Asp Arg Glu Leu His Arg Gln Thr Gly	
1230 1235 1240	
AAT GGA TTC CGC TGC CTC AAC TGC AAT GAC AAC ACT GAT GGC ATT CAC	309
Asn Gly Phe Arg Cys Leu Asn Cys Asn Asp Asn Thr Asp Gly Ile His	
1245 1250 1255	
TGC GAG AAG TGC AAG AAT GGC TTT TAC CGG CAC AGA GAA AGG GAC CGC	357
Cys Glu Lys Cys Lys Asn Gly Phe Tyr Arg His Arg Glu Arg Asp Arg	
1260 1265 1270	
TGT TTG CCC TGC AAT TGT AAC TCC AAA GGT TCT CTT AGT GCT CGA TGT	405
Cys Leu Pro Cys Asn Cys Asn Ser Lys Gly Ser Leu Ser Ala Arg Cys	
1275 1280 1285 1290	
GAC AAC TCT GGA CGG TGC AGC TGT AAA CCA GGT GTG ACA GGA GCC AGA	453

Asp	Asn	Ser	Gly	Arg	Cys	Ser	Cys	Lys	Pro	Gly	Val	Thr	Gly	Ala	Arg		
				1295				1300				1305					
TGC	GAC	CGA	TGT	CTG	CCA	GGC	TTC	CAC	ATG	CTC	ACG	GAT	GCG	GGG	TGC	501	
Cys	Asp	Arg	Cys	Leu	Pro	Gly	Phe	His	Met	Leu	Thr	Asp	Ala	Gly	Cys		
				1310				1315				1320					
ACC	CAA	GAC	CAG	AGA	CTG	CTA	GAC	TCC	AAG	TGT	GAC	TGT	GAC	CCA	GCT	549	
Thr	Gln	Asp	Gln	Arg	Leu	Leu	Asp	Ser	Lys	Cys	Asp	Cys	Asp	Pro	Ala		
				1325				1330				1335					
GGC	ATC	GCA	GGG	CCC	TGT	GAC	GCG	GGC	CGC	TGT	GTC	TGC	AAG	CCA	GCT	597	
Gly	Ile	Ala	Gly	Pro	Cys	Asp	Ala	Gly	Arg	Cys	Val	Cys	Lys	Pro	Ala		
				1340				1345				1350					
GTT	ACT	GGA	GAA	CGC	TGT	GAT	AGG	TGT	CGA	TCA	GGT	TAC	TAT	AAT	CTG	645	
Val	Thr	Gly	Glu	Arg	Cys	Asp	Arg	Cys	Arg	Ser	Gly	Tyr	Tyr	Asn	Leu		
				1355				1360				1365				1370	
GAT	GGG	GGG	AAC	CCT	GAG	GGC	TGT	ACC	CAG	TGT	TTC	TGC	TAT	GGG	CAT	693	
Asp	Gly	Gly	Asn	Pro	Glu	Gly	Cys	Thr	Gln	Cys	Phe	Cys	Tyr	Gly	His		
				1375				1380				1385					
TCA	GCC	AGC	TGC	CGC	AGC	TCT	GCA	GAA	TAC	AGT	GTC	CAT	AAG	ATC	ACC	741	
Ser	Ala	Ser	Cys	Arg	Ser	Ser	Ala	Glu	Tyr	Ser	Val	His	Lys	Ile	Thr		
				1390				1395				1400					
TCT	ACC	TTT	CAT	CAA	GAT	GTT	GAT	GGC	TGG	AAG	GCT	GTC	CAA	CGA	AAT	789	
Ser	Thr	Phe	His	Gln	Asp	Val	Asp	Gly	Trp	Lys	Ala	Val	Gln	Arg	Asn		
				1405				1410				1415					
GGG	TCT	CCT	GCA	AAG	CTC	CAA	TGG	TCA	CAG	CGC	CAT	CAA	GAT	GTG	TTT	837	
Gly	Ser	Pro	Ala	Lys	Leu	Gln	Trp	Ser	Gln	Arg	His	Gln	Asp	Val	Phe		
				1420				1425				1430					
AGC	TCA	GCC	CAA	CGA	CTA	GAT	CCT	GTC	TAT	TTT	GTG	GCT	CCT	GCC	AAA	885	
Ser	Ser	Ala	Gln	Arg	Leu	Asp	Pro	Val	Tyr	Phe	Val	Ala	Pro	Ala	Lys		
				1435				1440				1445				1450	
TTT	CTT	GGG	AAT	CAA	CAG	GTG	AGC	TAT	GGG	CAA	AGC	CTG	TCC	TTT	GAC	933	
Phe	Leu	Gly	Asn	Gln	Gln	Val	Ser	Tyr	Gly	Gln	Ser	Leu	Ser	Phe	Asp		
				1455				1460				1465					
TAC	CGT	GTG	GAC	AGA	GGA	GGC	AGA	CAC	CCA	TCT	GCC	CAT	GAT	GTG	ATC	981	
Tyr	Arg	Val	Asp	Arg	Gly	Gly	Arg	His	Pro	Ser	Ala	His	Asp	Val	Ile		
				1470				1475				1480					
CTG	GAA	GGT	GCT	GGT	CTA	CGG	ATC	ACA	GCT	CCC	TTG	ATG	CCA	CTT	GGC	1029	
Leu	Glu	Gly	Ala	Gly	Leu	Arg	Ile	Thr	Ala	Pro	Leu	Met	Pro	Leu	Gly		
				1485				1490				1495					
AAG	ACA	CTG	CCT	TGT	GGG	CTC	ACC	AAG	ACT	TAC	ACA	TTC	AGG	TTA	AAT	1077	
Lys	Thr	Leu	Pro	Cys	Gly	Leu	Thr	Lys	Thr	Tyr	Thr	Phe	Arg	Leu	Asn		
				1500				1505				1510					
GAG	CAT	CCA	AGC	AAT	AAT	TGG	AGC	CCC	CAG	CTG	AGT	TAC	TTT	GAG	TAT	1125	
Glu	His	Pro	Ser	Asn	Asn	Trp	Ser	Pro	Gln	Leu	Ser	Tyr	Phe	Glu	Tyr		
				1515				1520				1525				1530	
CGA	AGG	TTA	CTG	CGG	AAT	CTC	ACA	GCC	CTC	CGC	ATC	CGA	GCT</				





Pro	Met	Gly	Ser	Glu	Pro	Val	Gly	Cys	Arg	Ser	Asp	Gly	Thr	Cys	Val		
				1775				1780				1785					
TGC	AAG	CCA	GGA	TTT	GGT	GGC	CCC	AAC	TGT	GAG	CAT	GGA	GCA	TTC	AGC	1941	
Cys	Lys	Pro	Gly	Phe	Gly	Gly	Pro	Asn	Cys	Glu	His	Gly	Ala	Phe	Ser		
				1790				1795				1800					
TGT	CCA	GCT	TGC	TAT	AAT	CAA	GTG	AAG	ATT	CAG	ATG	GAT	CAG	TTT	ATG	1989	
Cys	Pro	Ala	Cys	Tyr	Asn	Gln	Val	Lys	Ile	Gln	Met	Asp	Gln	Phe	Met		
				1805				1810				1815					
CAG	CAG	CTT	CAG	AGA	ATG	GAG	GCC	CTG	ATT	TCA	AAG	GCT	CAG	GGT	GGT	2037	
Gln	Gln	Leu	Gln	Arg	Met	Glu	Ala	Leu	Ile	Ser	Lys	Ala	Gln	Gly	Gly		
				1820				1825				1830					
GAT	GGA	GTA	GTA	CCT	GAT	ACA	GAG	CTG	GAA	GGC	AGG	ATG	CAG	CAG	GCT	2085	
Asp	Gly	Val	Val	Pro	Asp	Thr	Glu	Leu	Glu	Gly	Arg	Met	Gln	Gln	Ala		
				1835				1840				1845				1850	
GAG	CAG	GCC	CTT	CAG	GAC	ATT	CTG	AGA	GAT	GCC	CAG	ATT	TCA	GAA	GGT	2133	
Glu	Gln	Ala	Leu	Gln	Asp	Ile	Leu	Arg	Asp	Ala	Gln	Ile	Ser	Glu	Gly		
				1855				1860				1865					
GCT	AGC	AGA	TCC	CTT	GGT	CTC	CAG	TTG	GCC	AAG	GTG	AGG	AGC	CAA	GAG	2181	
Ala	Ser	Arg	Ser	Leu	Gly	Leu	Gln	Leu	Ala	Lys	Val	Arg	Ser	Gln	Glu		
				1870				1875				1880					
AAC	AGC	TAC	CAG	AGC	CGC	CTG	GAT	GAC	CTC	AAG	ATG	ACT	GTG	GAA	AGA	2229	
Asn	Ser	Tyr	Gln	Ser	Arg	Leu	Asp	Asp	Leu	Lys	Met	Thr	Val	Glu	Arg		
				1885				1890				1895					
GTT	CGG	GCT	CTG	GGA	AGT	CAG	TAC	CAG	AAC	CGA	GTT	CGG	GAT	ACT	CAC	2277	
Val	Arg	Ala	Leu	Gly	Ser	Gln	Tyr	Gln	Asn	Arg	Val	Arg	Asp	Thr	His		
				1900				1905				1910					
AGG	CTC	ATC	ACT	CAG	ATG	CAG	CTG	AGC	CTG	GCA	GAA	AGT	GAA	GCT	TCC	2325	
Arg	Leu	Ile	Thr	Gln	Met	Gln	Leu	Ser	Leu	Ala	Glu	Ser	Glu	Ala	Ser		
				1915				1920				1925				1930	
TTG	GGA	AAC	ACT	AAC	ATT	CCT	GCC	TCA	GAC	CAC	TAC	GTG	GGG	CCA	AAT	2373	
Leu	Gly	Asn	Thr	Asn	Ile	Pro	Ala	Ser	Asp	His	Tyr	Val	Gly	Pro	Asn		
				1935				1940				1945					
GGC	TTT	AAA	AGT	CTG	GCT	CAG	GAG	GCC	ACA	AGA	TTA	GCA	GAA	AGC	CAC	2421	
Gly	Phe	Lys	Ser	Leu	Ala	Gln	Glu	Ala	Thr	Arg	Leu	Ala	Glu	Ser	His		
				1950				1955				1960					
GTT	GAG	TCA	GCC	AGT	AAC	ATG	GAG	CAA	CTG	ACA	AGG	GAA	ACT	GAG	GAC	2469	
Val	Glu	Ser	Ala	Ser	Asn	Met	Glu	Gln	Leu	Thr	Arg	Glu	Thr	Glu	Asp		
				1965				1970				1975					
TAT	TCC	AAA	CAA	GCC	CTC	TCA	CTG	GTG	CGC	AAG	GCC	CTG	CAT	GAA	GGA	2517	
Tyr	Ser	Lys	Gln	Ala	Leu	Ser	Leu	Val	Arg	Lys	Ala	Leu	His	Glu	Gly		
				1980				1985				1990					
GTC	GGA	AGC	GGA	AGC	GGT	AGC	CCG	GAC	GGT	GCT	GTG	GTG	CAA	GGG	CTT	2565	
Val	Gly	Ser	Gly	Ser	Gly	Ser	Pro	Asp	Gly	Ala	Val	Val	Gln	Gly	Leu		
				1995				2000				2005				2010	
GTG	GAA	AAA	TTG	GAG	AAA	ACC	AAG	TCC	CTG	GCC	CAG	CAG</					

Val	Glu	Lys	Leu	Glu	Lys	Thr	Lys	Ser	Leu	Ala	Gln	Gln	Leu	Thr	Arg		
				2015				2020				2025					
GAG Glu	GCC Ala	ACT Thr	CAA Gln	GCG Ala	GAA Glu	ATT Ile	GAA Glu	GCA Ala	GAT Asp	AGG Arg	TCT Ser	TAT Tyr	CAG Gln	CAC His	AGT Ser	2661	
				2030				2035				2040					
CTC Leu	CGC Arg	CTC Leu	CTG Leu	GAT Asp	TCA Ser	GTG Val	TCT Ser	CCG Pro	CTT Leu	CAG Gln	GGA Gly	GTC Val	AGT Ser	GAT Asp	CAG Gln	2709	
				2045				2050				2055					
TCC Ser	TTT Phe	CAG Gln	GTG Val	GAA Glu	GAA Glu	GCA Ala	AAG Lys	AGG Arg	ATC Ile	AAA Lys	CAA Gln	AAA Lys	GCG Ala	GAT Asp	TCA Ser	2757	
				2060				2065				2070					
CTC Leu	TCA Ser	AGC Ser	CTG Leu	GTA Val	ACC Thr	AGG Arg	CAT His	ATG Met	GAT Asp	GAG Glu	TTC Phe	AAG Lys	CGT Arg	ACA Thr	CAA Gln	2805	
				2080				2085				2090					
AAG Lys	AAT Asn	CTG Leu	GGA Gly	AAC Asn	TGG Trp	AAA Lys	GAA Glu	GAA Glu	GCA Ala	CAG Gln	CAG Gln	CTC Leu	TTA Leu	CAG Gln	AAT Asn	2853	
				2095				2100				2105					
GGA Gly	AAA Lys	AGT Ser	GGG Gly	AGA Arg	GAG Glu	AAA Lys	TCA Ser	GAT Asp	CAG Gln	CTG Leu	CTT Leu	TCC Ser	CGT Arg	GCC Ala	AAT Asn	2901	
				2110				2115				2120					
CTT Leu	GCT Ala	AAA Lys	AGC Ser	AGA Arg	GCA Ala	CAA Gln	GAA Glu	GCA Ala	CTG Leu	AGT Ser	ATG Met	GGC Gly	AAT Asn	GCC Ala	ACT Thr	2949	
				2125				2130				2135					
TTT Phe	TAT Tyr	GAA Glu	GTT Val	GAG Glu	AGC Ser	ATC Ile	CTT Leu	AAA Lys	AAC Asn	CTC Leu	AGA Arg	GAG Glu	TTT Phe	GAC Asp	CTG Leu	2997	
				2140				2145				2150					
CAG Gln	GTG Val	GAC Asp	AAC Asn	AGA Arg	AAA Lys	GCA Ala	GAA Glu	GCT Ala	GAA Glu	GAA Glu	GCC Ala	ATG Met	AAG Lys	AGA Arg	CTC Leu	3045	
				2155				2160				2165				2170	
TCC Ser	TAC Tyr	ATC Ile	AGC Ser	CAG Gln	AAG Lys	GTT Val	TCA Ser	GAT Asp	GCC Ala	AGT Ser	GAC Asp	AAG Lys	ACC Thr	CAG Gln	CAA Gln	3093	
				2175				2180				2185					
GCA Ala	GAA Glu	AGA Arg	GCC Ala	CTG Leu	GGG Gly	AGC Ser	GCT Ala	GCT Ala	GCT Ala	GAT Asp	GCA Ala	CAG Gln	AGG Arg	GCA Ala	AAG Lys	3141	
				2190				2195				2200					
AAT Asn	GGG Gly	GCC Ala	GGG Gly	GAG Glu	GCC Ala	CTG Leu	GAA Glu	ATC Ile	TCC Ser	AGT Ser	GAG Glu	ATT Ile	GAA Glu	CAG Gln	GAG Glu	3189	
				2205				2210				2215					
ATT Ile	GGG Gly	AGT Ser	CTG Leu	AAC Asn	TTG Leu	GAA Glu	GCC Ala	AAT Asn	GTG Val	ACA Thr	GCA Ala	GAT Asp	GGA Gly	GCC Ala	TTG Leu	3237	
				2220				2225				2230					
GCC Ala	ATG Met	GAA Glu	AAG Lys	GGA Gly	CTG Leu	GCC Ala	TCT Ser	CTG Leu	AAG Lys	AGT Ser	GAG Glu	ATG Met	AGG Arg	GAA Glu	GTG Val	3285	
				2235				2240				2245				2250	
GAA	GGA	GAG	CTG	GAA	AGG	AAG	GAG	CTG	GAG	TTT	GAC	ACG	AAT	ATG	GAT	3333	

1112 amino acids, linear topology, protein molecule

```
Glu Gly Glu Leu Glu Arg Lys Glu Leu Glu Phe Asp Thr Asn Met Asp
      2255                      2260                      2265

GCA GTA CAG ATG GTG ATT ACA GAA GCC CAG AAG GTT GAT ACC AGA GCC      3381
Ala Val Gln Met Val Ile Thr Glu Ala Gln Lys Val Asp Thr Arg Ala
      2270                      2275                      2280

AAG AAC GCT GGG GTT ACA ATC CAA GAC ACA CTC AAC ACA TTA GAC GGC      3429
Lys Asn Ala Gly Val Thr Ile Gln Asp Thr Leu Asn Thr Leu Asp Gly
      2285                      2290                      2295

CTC CTG CAT CTG ATG GGT ATG TGA ACCCACAACC CACAACCTTC CAGCTCCATG      3483
Leu Leu His Leu Met Gly Met *
      2300                      2305

CTCCAGGGCT TTGCTCCAGA ACACTCACTA TACCTAGCCC CAGCAAAGGG GAGTCTCAGC      3543

TTTCCTTAAG GATATCAGTA AATGTGCTTT GTTTCAGGC CCAGATAACT TTCGGCAGGT      3603

TCCCTTACAT TTA CTG GACC CTGTTTACC GTTGCTAAGA TGGGTCAC TG AACACCTATT      3663

GCACTTGGGG GTAAAGGTCT GTGGGCCAAA GAACAGGTGT ATATAAGCAA CTTACAGAA      3723

CACGAGACAG CTTGGGAATC CTGCTAAAGA GTCTGGCCTG GACCCTGAGA AGCCAGTGGA      3783

CAGTTTAAAG CAGAGGAATA ACATCACCAC TGTATATTTC AGAAAGATCA CTAGGGCAGC      3843

CGAGTGGAGG AAAGCTTGAA GAGGGGGTTA GAGAGAAGGC AGGTTGAGAC TACTTAAGAT      3903

ATTGTTGAAA TAATTGAAGA GAGAAATGAC AGGAGCCTGC TCTAAGGCAG TAGAATGGTG      3963

GCTGGGAAGA TGTGAAGGAA GATTTTCCCA GTCTGTGAAG TCAAGAATCA CTTGCCGGCC      4023

GGGTGTGGTG GCTCACGCCT GTAATTCTAG CACTTTGGGA GACTGAAGCG GGTGGATCAC      4083

CCGAGGTCAG GAGTTGAAGA CCAGCCTGGC CAACATGGTG AAACCCTGTC TCTACTAAAA      4143

GTACAAAAAT TAGCTGGATG ATGGTGGTGG GCGCCTGTAA TTCCAGCTAC TCAGGAGTCT      4203

GAGGCAGGAG AATCGCTTGA ACCCAGGAGG CGAGGTTACA GTGAGCCAAG ATTGCACCAC      4263

TGCTCTTCCA GCCTGGGAAC AGAGAGACTG CCTAAAAAAA AAAAAAAAAA AAA      4316
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1112 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Pro Ala Leu Trp Leu Gly Cys Cys Leu Cys Phe Ser Leu Leu Leu
  1             5             10             15

Pro Ala Ala Arg Ala Thr Ser Arg Arg Glu Val Cys Asp Cys Asn Gly
      20             25             30
```

1000 900 800 700 600 500 400 300 200 100 0

Lys Ser Arg Gln Cys Ile Phe Asp Arg Glu Leu His Arg Gln Thr Gly  
35 40 45

Asn Gly Phe Arg Cys Leu Asn Cys Asn Asp Asn Thr Asp Gly Ile His  
50 55 60

Cys Glu Lys Cys Lys Asn Gly Phe Tyr Arg His Arg Glu Arg Asp Arg  
65 70 75 80

Cys Leu Pro Cys Asn Cys Asn Ser Lys Gly Ser Leu Ser Ala Arg Cys  
85 90 95

Asp Asn Ser Gly Arg Cys Ser Cys Lys Pro Gly Val Thr Gly Ala Arg  
100 105 110

Cys Asp Arg Cys Leu Pro Gly Phe His Met Leu Thr Asp Ala Gly Cys  
115 120 125

Thr Gln Asp Gln Arg Leu Leu Asp Ser Lys Cys Asp Cys Asp Pro Ala  
130 135 140

Gly Ile Ala Gly Pro Cys Asp Ala Gly Arg Cys Val Cys Lys Pro Ala  
145 150 155 160

Val Thr Gly Glu Arg Cys Asp Arg Cys Arg Ser Gly Tyr Tyr Asn Leu  
165 170 175

Asp Gly Gly Asn Pro Glu Gly Cys Thr Gln Cys Phe Cys Tyr Gly His  
180 185 190

Ser Ala Ser Cys Arg Ser Ser Ala Glu Tyr Ser Val His Lys Ile Thr  
195 200 205

Ser Thr Phe His Gln Asp Val Asp Gly Trp Lys Ala Val Gln Arg Asn  
210 215 220

Gly Ser Pro Ala Lys Leu Gln Trp Ser Gln Arg His Gln Asp Val Phe  
225 230 235 240

Ser Ser Ala Gln Arg Leu Asp Pro Val Tyr Phe Val Ala Pro Ala Lys  
245 250 255

Phe Leu Gly Asn Gln Gln Val Ser Tyr Gly Gln Ser Leu Ser Phe Asp  
260 265 270

Tyr Arg Val Asp Arg Gly Gly Arg His Pro Ser Ala His Asp Val Ile  
275 280 285

Leu Glu Gly Ala Gly Leu Arg Ile Thr Ala Pro Leu Met Pro Leu Gly  
290 295 300

Lys Thr Leu Pro Cys Gly Leu Thr Lys Thr Tyr Thr Phe Arg Leu Asn  
305 310 315 320

Glu His Pro Ser Asn Asn Trp Ser Pro Gln Leu Ser Tyr Phe Glu Tyr  
325 330 335

Arg Arg Leu Leu Arg Asn Leu Thr Ala Leu Arg Ile Arg Ala Thr Tyr  
340 345 350



Ala	Ser	Arg	Ser	Leu	Gly	Leu	Gln	Leu	Ala	Lys	Val	Arg	Ser	Gln	Glu	
	675						680					685				
Asn	Ser	Tyr	Gln	Ser	Arg	Leu	Asp	Asp	Leu	Lys	Met	Thr	Val	Glu	Arg	
	690					695					700					
Val	Arg	Ala	Leu	Gly	Ser	Gln	Tyr	Gln	Asn	Arg	Val	Arg	Asp	Thr	His	
705					710					715					720	
Arg	Leu	Ile	Thr	Gln	Met	Gln	Leu	Ser	Leu	Ala	Glu	Ser	Glu	Ala	Ser	
				725					730					735		
Leu	Gly	Asn	Thr	Asn	Ile	Pro	Ala	Ser	Asp	His	Tyr	Val	Gly	Pro	Asn	
			740					745					750			
Gly	Phe	Lys	Ser	Leu	Ala	Gln	Glu	Ala	Thr	Arg	Leu	Ala	Glu	Ser	His	
		755					760					765				
Val	Glu	Ser	Ala	Ser	Asn	Met	Glu	Gln	Leu	Thr	Arg	Glu	Thr	Glu	Asp	
	770					775					780					
Tyr	Ser	Lys	Gln	Ala	Leu	Ser	Leu	Val	Arg	Lys	Ala	Leu	His	Glu	Gly	
785					790					795					800	
Val	Gly	Ser	Gly	Ser	Gly	Ser	Pro	Asp	Gly	Ala	Val	Val	Gln	Gly	Leu	
				805					810					815		
Val	Glu	Lys	Leu	Glu	Lys	Thr	Lys	Ser	Leu	Ala	Gln	Gln	Leu	Thr	Arg	
			820					825					830			
Glu	Ala	Thr	Gln	Ala	Glu	Ile	Glu	Ala	Asp	Arg	Ser	Tyr	Gln	His	Ser	
		835					840					845				
Leu	Arg	Leu	Leu	Asp	Ser	Val	Ser	Pro	Leu	Gln	Gly	Val	Ser	Asp	Gln	
	850					855					860					
Ser	Phe	Gln	Val	Glu	Glu	Ala	Lys	Arg	Ile	Lys	Gln	Lys	Ala	Asp	Ser	
865					870					875					880	
Leu	Ser	Ser	Leu	Val	Thr	Arg	His	Met	Asp	Glu	Phe	Lys	Arg	Thr	Gln	
			885						890					895		
Lys	Asn	Leu	Gly	Asn	Trp	Lys	Glu	Glu	Ala	Gln	Gln	Leu	Leu	Gln	Asn	
		900						905						910		
Gly	Lys	Ser	Gly	Arg	Glu	Lys	Ser	Asp	Gln	Leu	Leu	Ser	Arg	Ala	Asn	
		915					920					925				
Leu	Ala	Lys	Ser	Arg	Ala	Gln	Glu	Ala	Leu	Ser	Met	Gly	Asn	Ala	Thr	
	930					935					940					
Phe	Tyr	Glu	Val	Glu	Ser	Ile	Leu	Lys	Asn	Leu	Arg	Glu	Phe	Asp	Leu	
945					950					955					960	
Gln	Val	Asp	Asn	Arg	Lys	Ala	Glu	Ala	Glu	Glu	Ala	Met	Lys	Arg	Leu	
			965						970					975		
Ser	Tyr	Ile	Ser	Gln	Lys	Val	Ser	Asp	Ala	Ser	Asp	Lys	Thr	Gln	Gln	
		980						985					990			

1000  
 900  
 800  
 700  
 600  
 500  
 400  
 300  
 200  
 100  
 0

Ala	Glu	Arg	Ala	Leu	Gly	Ser	Ala	Ala	Ala	Asp	Ala	Gln	Arg	Ala	Lys
							995			1000					1005
Asn	Gly	Ala	Gly	Glu	Ala	Leu	Glu	Ile	Ser	Ser	Glu	Ile	Glu	Gln	Glu
							1010								1020
Ile	Gly	Ser	Leu	Asn	Leu	Glu	Ala	Asn	Val	Thr	Ala	Asp	Gly	Ala	Leu
							1025								1040
Ala	Met	Glu	Lys	Gly	Leu	Ala	Ser	Leu	Lys	Ser	Glu	Met	Arg	Glu	Val
							1045								1055
Glu	Gly	Glu	Leu	Glu	Arg	Lys	Glu	Leu	Glu	Phe	Asp	Thr	Asn	Met	Asp
							1060								1070
Ala	Val	Gln	Met	Val	Ile	Thr	Glu	Ala	Gln	Lys	Val	Asp	Thr	Arg	Ala
							1075								1085
Lys	Asn	Ala	Gly	Val	Thr	Ile	Gln	Asp	Thr	Leu	Asn	Thr	Leu	Asp	Gly
							1090								1100
Leu	Leu	His	Leu	Met	Gly	Met									*
							1105								1110

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligomer primers"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAGCGCAGAG TGAGAACCAC

20

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligomer primers"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACTGTATTCT GCAGAGCTGC

20



(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
    (A) DESCRIPTION: /desc = "oligomer primers"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTCCTTTCCCTACCTTGTG

20

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
    (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGTGGAAGCCTGGCAGACAT

20

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 720 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
    (A) DESCRIPTION: /desc = "OLIGOMER PRIMER"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Gly Thr Cys Thr Thr Thr Ala Thr Ala Gly Gly Gly Ala Gly Gly  
1                    5                    10                    15

Thr Thr Gly Gly Cys Cys Ala Gly Thr Cys Ala Ala Thr Ala Gly Gly  
                    20                    25                    30

Thr Thr Ala Cys Thr Thr Thr Ala Thr Gly Ala Gly Thr Thr Gly Cys  
                    35                    40                    45

Thr Ala Ala Cys Cys Cys Thr Gly Gly Thr Gly Ala Gly Cys Ala Gly  
                    50                    55                    60

Gly Ala Ala Gly Thr Thr Ala Thr Gly Thr Gly Gly Ala Cys Cys Ala  
65                    70                    75                    80

Gly Gly Ala Gly Ala Gly Ala Ala Cys Cys Cys Thr Thr Gly Gly

85

90

95

Thr	Thr	Cys	Ala	Gly	Cys	Cys	Thr	Gly	Gly	Ala	Gly	Ala	Ala	Ala	Gly
			100					105					110		
Gly	Ala	Gly	Ala	Gly	Gly	Thr	Thr	Gly	Ala	Cys	Cys	Cys	Thr	Ala	Ala
		115					120					125			
Ala	Cys	Thr	Gly	Gly	Ala	Gly	Gly	Gly	Thr	Gly	Gly	Ala	Gly	Ala	Gly
	130					135					140				
Gly	Ala	Cys	Cys	Cys	Thr	Gly	Thr	Thr	Gly	Thr	Gly	Ala	Cys	Thr	Cys
145					150					155					160
Thr	Cys	Cys	Gly	Ala	Cys	Thr	Gly	Ala	Cys	Thr	Thr	Gly	Thr	Cys	Thr
			165						170					175	
Thr	Cys	Cys	Thr	Thr	Gly	Ala	Thr	Gly	Thr	Cys	Cys	Thr	Thr	Thr	Ala
			180						185					190	
Ala	Gly	Cys	Cys	Gly	Gly	Ala	Gly	Cys	Thr	Gly	Ala	Thr	Thr	Cys	Gly
		195					200					205			
Gly	Gly	Cys	Thr	Gly	Cys	Thr	Gly	Cys	Cys	Thr	Thr	Ala	Thr	Thr	Thr
	210					215						220			
Cys	Thr	Gly	Ala	Gly	Thr	Thr	Ala	Gly	Cys	Gly	Cys	Thr	Cys	Thr	Thr
225					230					235					240
Ala	Ala	Gly	Ala	Thr	Thr	Gly	Gly	Gly	Cys	Cys	Thr	Cys	Cys	Cys	Ala
				245					250					255	
Gly	Thr	Thr	Thr	Gly	Ala	Gly	Gly	Ala	Ala	Gly	Gly	Gly	Gly	Cys	Gly
			260					265						270	
Gly	Gly	Cys	Thr	Gly	Cys	Thr	Gly	Thr	Cys	Thr	Ala	Cys	Cys	Thr	Cys
		275					280					285			
Thr	Gly	Thr	Gly	Ala	Ala	Thr	Cys	Thr	Gly	Cys	Cys	Cys	Thr	Gly	Gly
	290					295					300				
Ala	Cys	Cys	Ala	Cys	Cys	Cys	Cys	Gly	Gly	Gly	Ala	Gly	Ala	Gly	Ala
305					310					315					320
Ala	Gly	Gly	Ala	Gly	Gly	Gly	Cys	Thr	Cys	Cys	Gly	Gly	Gly	Gly	Ala
				325					330					335	
Ala	Thr	Cys	Thr	Cys	Gly	Cys	Ala	Cys	Ala	Thr	Thr	Cys	Cys	Ala	Gly
			340					345						350	
Gly	Cys	Ala	Ala	Ala	Gly	Gly	Cys	Thr	Cys	Cys	Cys	Gly	Gly	Gly	Cys
		355					360					365			
Cys	Gly	Cys	Ala	Gly	Cys	Cys	Thr	Cys	Thr	Gly	Thr	Gly	Cys	Cys	Ala
	370					375					380				
Cys	Ala	Cys	Cys	Cys	Thr	Thr	Gly	Gly	Cys	Cys	Cys	Gly	Gly	Gly	Cys
385					390					395					400
Cys	Ala	Gly	Gly	Thr	Gly	Thr	Gly	Cys	Gly	Cys	Cys	Cys	Thr	Cys	Cys

405

410

415

Thr Cys Gly Cys Thr Gly Cys Gly Ala Gly Gly Gly Gly Gly Ala Gly  
 420 425 430

Cys Gly Gly Gly Cys Gly Gly Cys Thr Gly Cys Gly Gly Gly Gly Ala  
 435 440 445

Gly Cys Gly Ala Thr Thr Thr Thr Cys Cys Ala Gly Cys Cys Cys Gly  
 450 455 460

Gly Thr Thr Thr Gly Thr Gly Cys Thr Cys Thr Gly Thr Gly Thr Gly  
 465 470 475 480

Thr Thr Thr Gly Thr Cys Thr Gly Cys Cys Thr Cys Thr Gly Gly Ala  
 485 490 495

Gly Gly Gly Cys Thr Gly Gly Gly Thr Cys Cys Thr Cys Cys Thr Thr  
 500 505 510

Ala Thr Thr Cys Ala Cys Ala Gly Gly Thr Gly Ala Gly Thr Cys Ala  
 515 520 525

Cys Ala Cys Cys Cys Thr Gly Ala Ala Ala Cys Ala Cys Ala Gly Gly  
 530 535 540

Cys Thr Cys Thr Cys Thr Thr Cys Cys Thr Gly Thr Cys Ala Gly Gly  
 545 550 555 560

Ala Cys Thr Gly Ala Gly Thr Cys Ala Gly Gly Thr Ala Gly Ala Ala  
 565 570 575

Gly Ala Gly Thr Cys Gly Ala Thr Ala Ala Ala Ala Cys Cys Ala Cys  
 580 585 590

Cys Thr Gly Ala Thr Cys Ala Ala Gly Gly Ala Ala Ala Gly Gly  
 595 600 605

Ala Ala Gly Gly Cys Ala Cys Ala Gly Cys Gly Gly Ala Gly Cys Gly  
 610 615 620

Cys Ala Gly Ala Gly Thr Gly Ala Gly Ala Ala Cys Cys Ala Cys Cys  
 625 630 635 640

Ala Ala Cys Cys Gly Ala Gly Gly Cys Gly Cys Cys Gly Gly Gly Cys  
 645 650 655

Ala Gly Cys Gly Ala Cys Cys Cys Cys Thr Gly Cys Ala Gly Cys Gly  
 660 665 670

Gly Ala Gly Ala Cys Ala Gly Ala Gly Ala Cys Thr Gly Ala Gly Cys  
 675 680 685

Gly Gly Cys Cys Cys Gly Gly Cys Ala Cys Cys Gly Cys Cys Ala Thr  
 690 695 700

Gly Cys Cys Thr Gly Cys Gly Cys Thr Cys Thr Gly Gly Cys Thr Gly  
 705 710 715 720

1000 900 800 700 600 500 400 300 200 100 0